C1 cancel

acids of the disclosed human-*P. carinii MSG*s; these may be used for amplification or as probes. The sequences of these conserved nucleic acid molecule regions include residues 2794-3042 of *HMSGp1* (SEQ ID NO: 1), 2758-3006 of *HMSGp3* (SEQ ID NO: 3), 2845-3090 of *HMSG11* (SEQ ID NO: 5), 2839-3084 of *HMSG14* (SEQ ID NO: 7), 2836-3081 of *HMSG32* (SEQ ID NO: 9), 2809-3054 of *HMSG33* (SEQ ID NO: 11), 2821-3072 of *HMSG35* (SEQ ID NO: 13), or 1-249 of *HMSGp2* (SEQ ID NO: 15). In addition, this invention encompasses sequences with at least 70% sequence identity to these regions, and recombinant vectors comprising such nucleic acid molecules and conserved regions from within such nucleic acid molecules, as well as transgenic cells including such a recombinant vector.

Please replace the paragraph at page 3, lines 14-31 with the following:

C2

Another aspect of this invention provides a method of detecting the presence of *Pneumocystis carinii* in a biological specimen, by amplifying with a nucleic acid amplification method (*e.g.*, the polymerase chain reaction) a human-*P. carinii* nucleic acid sequence using two or more oligonucleotide primers derived from a human-*P. carinii* MSG protein encoding sequence, then determining whether an amplified sequence is present. In a preferred embodiment of this invention, the human-*P. carinii* nucleic acid sequence is a highly conserved region within an MSG-protein encoding sequence. Such a highly conserved region may, for instance, include residues 2794-3042 of *HMSGp1* (SEQ ID NO: 1), 2758-3006 of *HMSGp3* (SEQ ID NO: 3), 2845-3090 of *HMSG11* (SEQ ID NO: 5), 2839-3084 of *HMSG14* (SEQ ID NO: 7), 2836-3081 of *HMSG32* (SEQ ID NO: 9), 2809-3054 of *HMSG33* (SEQ ID NO: 11), 2821-3072 of *HMSG35* (SEQ ID NO: 13), or 1-249 of *HMSGp2* (SEQ ID NO: 15). A further aspect of this invention is the method of detecting the presence of *Pneumocystis carinii* in a biological specimen, by determining whether an amplified sequence is present, for instance by electrophoresis and staining of the amplified sequence, or hybridization to a labeled probe of the amplified sequence. Appropriate labels for the hybridization probe include a fluorescent molecule, a chemiluminescent molecule, an enzyme, a co-factor, an enzyme substrate, or a hapten. The nucleotide sequence of such a probe can be chosen from any *MSG* gene sequence that is amplified in the detection method, and for instance can include a nucleic acid sequence according to SEQ ID NO: 19.

Please replace the paragraph at page 8, lines 28-33 with the following:

C3 Further nucleic acid molecules might comprise at least 15 consecutive nucleotides of the regions encoding the conserved carboxy-terminal portion of each human-*P. carinii MSG* gene. These regions comprise nucleotides 2794-3042 of *HMSGp1* (SEQ ID NO: 1), 2758-3006 of *HMSGp3* (SEQ ID NO: 3), 2845-3090 of *HMSG11* (SEQ ID NO: 5), 2839-3084 of *HMSG14* (SEQ ID NO: 7), 2836-3081 of *HMSG32* (SEQ ID NO: 9), 2809-3054 of *HMSG33* (SEQ ID NO: 11), 2821-3072 of *HMSG35* (SEQ ID NO: 13), and 1-249 of *HMSGp2* (SEQ ID NO: 15), respectively.

Please replace the paragraph at page 9, lines 22-32 with the following:

C4 The NCBI Basic Local Alignment Search Tool (BLAST) (Altschul *et al.* (1990) *J. Mol. Biol.* 215:403-410) is available from several sources, including the National Center for Biotechnology Information (NCBI, Bethesda, MD) and on the Internet, for use in connection with the sequence analysis programs blastp, blastn, blastx, tblastn and tblastx. It can be accessed at the NCBI online site under the "BLAST" heading. A description of how to determine sequence identity using this program is available at the NCBI online site under the "BLAST" heading and "BLAST overview" subheading. For comparisons of amino acid sequences of greater than about 30 amino acids, the Blast 2 sequences function is employed using the default BLOSUM62 matrix set to default parameters, (gap existence cost of 11, and a per residue gap cost of 1). When aligning short peptides (fewer than around 30 amino acids), the alignment should be performed using the Blast 2 sequences function, employing the PAM30 matrix set to default parameters (open gap 9, extension gap 1 penalties).

Please replace the paragraph at page 9, line 33 through page 10, line 6 with the following:

C5 Other members of the gene family of the disclosed human-*P. carinii* MSG proteins typically possess at least 60% sequence identity counted over full-length alignment with the amino acid sequence of human-*P. carinii* MSG using the NCBI Blast 2.0, gapped blastp set to default parameters. Sequence identity over the about 100 C-terminal amino acids will typically be higher than 60%, for instances about 63%. Proteins with even greater similarity to the reference sequence will show increasing percentage identities when assessed by this method, such as at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, or at least 98%

C5 cont'd

sequence identity. When less than the entire sequence is being compared for sequence identity, homologs will typically possess at least 75% sequence identity over short windows of 10-20 amino acids, and may possess sequence identities of at least 85% or at least 90% or 95% depending on their similarity to the reference sequence. Methods for determining sequence identity over such short windows are described at the NCBI online site under the "BLAST" heading and "Frequently Asked Questions" subheading.

Please replace the paragraph at page 12, line 35, through page 13, line 12 with the following:

C6

Oligonucleotides that are derived from the human-*P. carinii HMSGp1*, *HMSGp3*, *HMSG11*, *HMSG14*, *HMSG32*, *HMSG33*, and *HMSG35* gene sequences (SEQ ID NOS: 1, 3, 5, 7, 9, 11, and 13, respectively), as well as the fragment of *HMSGp2* disclosed (SEQ ID NO: 15), are encompassed within the scope of the present invention. Preferably, such oligonucleotide primers will comprise a sequence of at least 15-20 consecutive nucleotides of the relevant human-*P. carinii MSG* gene sequence. To enhance amplification specificity, oligonucleotide primers comprising at least 25, 30, 35, 40, 45 or 50 consecutive nucleotides of these sequences may also be used. These primers for instance may be obtained from any region of the disclosed sequences. By way of example, human-*P. carinii MSG* gene sequences may be apportioned into halves or quarters based on sequence length, and the isolated nucleic acid molecules may be derived from the first or second halves of the molecules, or any of the four quarters. In addition, primers may be specifically chosen from the conserved carboxy-terminal region of each *MSG* coding sequence. This region comprises nucleic acid residues 2794-3042 of *HMSGp1* (SEQ ID NO: 1), 2758-3006 of *HMSGp3* (SEQ ID NO: 3), 2845-3090 of *HMSG11* (SEQ ID NO: 5), 2839-3084 of *HMSG14* (SEQ ID NO: 7), 2836-3081 of *HMSG32* (SEQ ID NO: 9), 2809-3054 of *HMSG33* (SEQ ID NO: 11), 2821-3072 of *HMSG35* (SEQ ID NO: 13), and 1-249 of *HMSGp2* (SEQ ID NO: 15).

Please replace the paragraph at page 17, line 34, though page 18, line 6 with the following:

C7

The selection of PCR primers will be made according to the portions of the gene sequence that are to be amplified. For use in PCR detection of *P. carinii*, it is advantageous to choose primer-annealing sites that are highly conserved across many different members of the human-*P. carinii MSG* gene family. For instance, it is advantageous to choose primer sites from

C7 cont'd within the regions of human-*P. carinii* sequence displaying greater than 63% sequence identity across the disclosed family members, *e.g.*, that portion of the gene encoding the conserved carboxy-terminal region of the protein. The highly conserved carboxy-terminal regions of the disclosed genes are as follows: residues 2794-3042 of *HMSGp1* (SEQ ID NO: 1), 2758-3006 of *HMSGp3* (SEQ ID NO: 3), 2845-3090 of *HMSG11* (SEQ ID NO: 5), 2839-3084 of *HMSG14* (SEQ ID NO: 7), 2836-3081 of *HMSG32* (SEQ ID NO: 9), 2809-3054 of *HMSG33* (SEQ ID NO: 11), 2821-3072 of *HMSG35* (SEQ ID NO: 13), and 1-249 of *HMSGp2* (SEQ ID NO: 15).

Please replace the paragraph at page 18, lines 20-32 with the following:

C8 The presence of amplified human-*P. carinii MSG* sequences can be determined in any conventional manner, including electrophoresis and staining (for instance, with ethidium bromide) of the amplified sequence, or hybridization of a labeled probe to the amplified sequence. For general guidelines on such techniques, see *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor, New York (1989), and *Current Protocols in Molecular Biology*, Greene Publishing Associates and Wiley-Intersciences (1987). Hybridization probes appropriate for use in detection of amplified human-*P. carinii MSG* sequences are essentially equivalent to those described above for direct hybridization. The region of the gene that has been amplified will be important in choosing an appropriate probe; the detection probe should hybridize to a sequence that falls between the ends of the amplification primers such that the annealing site of the probe is amplified. By way of example, one appropriate oligonucleotide probe is JKK16 (SEQ ID NO: 19), which corresponds to residues of 2926-2950 of *HMSG33*. This probe could be used for detection of both full-length and carboxy-terminal amplified fragments of human-*P. carinii MSG* genes.

Please replace the paragraph at page 27, lines5-12 with the following:

C9 ***MSG* sequence:** For PCR amplification of human-*P. carinii MSG* in clinical samples, the upstream primer used was an equimolar mixture of JKK14 (SEQ ID NO: 17) (corresponding to the residues of 2809-2833 of *HMSG33,* which is also 2845-2869 of *hMSG11*) and JKK15 (SEQ ID NO: 18) (corresponding to the residues of 2836-2860 of *HMSG32*). The downstream primer used was JKK17 (SEQ ID NO: 20) (complementary to the conserved residues 3028-3052 of *HMSG33,* which is also 3064-3088 of *MSG11*). In experiments wherein the amplified product